

UNITED STATES DISTRICT COURT
DISTRICT OF MASSACHUSETTS

UNITED STATES OF AMERICA

v.

21-cr-10104 (PBS)

VLADISLAV KLYUSHIN,
a/k/a “Vladislav Kliushin”
IVAN ERMAKOV,
a/k/a “Ivan Yermakov,” and
NIKOLAI RUMIANTCEV,
a/k/a “Nikolay Rumyantsev,”

Defendants.

OPPOSITION OF THE UNITED STATES TO DEFENDANT’S SUPPLEMENTAL REQUEST
FOR DAUBERT HEARING AND TO EXCLUDE TESTIMONY OF MAXWELL CLARKE

In further support of his motion to exclude the testimony of the government’s expert witness, SEC Financial Economist Maxwell Clarke, the defendant makes three principal arguments: (1) that Mr. Clarke’s analysis rests on faulty data and assumptions, including that the stock market is a game of chance; (2) that Mr. Clarke improperly employed a statistical technique known as Fisher’s exact test; and (3) that the Court should exercise its discretion to exclude the analysis under Fed. R. Evid. 403. These contentions are without merit and should be rejected.

I. Klyushin’s Claims of Faulty Data and Assumptions are Meritless

Klyushin makes three principal claims of error in the data and assumptions on which Mr. Clarke relied for his analysis: first, that Mr. Clarke erred with respect to whether the earnings report of one particular company, Avnet, was handled by FA 1 in January 2020; second, that Mr. Clarke’s data is missing “earnings information available for some but not all transactions;” and third, that the data is faulty because Mr. Clarke did not use the client lists of FA 1 and FA 2 to attribute earnings

reports to them, excluded trading in companies unrelated to FA 1 and FA 2, and excluded trading that did not immediately precede an earnings announcement. These contentions are without merit.

A. Klyushin's Claim About Avnet is Misleading and Irrelevant to the Statistical Analysis

Klyushin first contends that Mr. Clarke erred because his data suggests that FA 1 was not associated with Avnet's January 2020 earnings report. If that is true, he argues, then Klyushin could not have traded on the basis of Avnet MNPI because FA 1 "did not have that information." Dkt 131 at 3. And if it is *not* true, he contends, then Mr. Clarke's entire data set is suspect. *Id.* Klyushin's assumptions are simply wrong. Mr. Clarke's data was not flawed, nor, in any event, would the single Avnet earnings report Klyushin identifies have any effect on Mr. Clarke's analysis.

As explained below, the evidence will show that Avnet filed its own earnings report in January 2020 using software provided by FA 1. Accordingly, while FA 1 itself did not file the report, the earnings report resided in FA 1's computer network and was downloaded from that network prior to the conspirators' trading. The fact that Avnet (and other issuers) self-filed from time to time does not undermine Mr. Clarke's analysis or affect the significance of the correlation he tested, which was between the conspirators' trading and whether or not earnings reports were filed by FA 1 and FA 2. At most, the fact that Klyushin also traded on an earnings report that Avnet self-filed, but that resided in FA 1's computer network, would provide *additional* evidence of his crime.

As noted in the government's prior filing, the SEC assigns a unique "accession number" to every filing in its EDGAR system. The accession number is "a unique identifier assigned automatically to an accepted submission by the EDGAR Filer System. The first 10 digits . . . comprise the Central Index Key (CIK) of the entity submitting the filing. This could be the company or a third-party filer

agent.”¹ For the period of the charged conspiracy—January 1, 2018 through September 30, 2020—Mr. Clarke determined the accession number of every earnings report filed with the SEC. He determined that, of the 38,359 earnings events during the period of Mr. Klyushin’s trading, 44 percent had accession numbers associated with FA 1 or FA 2.

Of the issuers who file their own earnings reports (rather than using a filing agent), the overwhelming majority use software sold by a company called Workiva. A small number of companies, however, use software offered by filing agents—much like a taxpayer might either hire H&R Block to prepare and file their taxes, or use H&R Block’s online system to prepare and file their own taxes. This was the case with Avnet. As the SEC’s EDGAR database makes clear, Avnet’s January 23, 2020 8-K filing was self-filed—hence, the first ten digits of the accession number (0000008858) are those of Avnet itself, *not* a filing agent.²

```
<SEC-DOCUMENT>0000008858-20-000007.txt : 20200123
<SEC-HEADER>0000008858-20-000007.hdr.sgml : 20200123
<ACCEPTANCE-DATETIME>20200123160751
ACCESSION NUMBER:      0000008858-20-000007
CONFORMED SUBMISSION TYPE: 8-K
PUBLIC DOCUMENT COUNT:  14
CONFORMED PERIOD OF REPORT: 20200123
ITEM INFORMATION:       Results of Operations and Financial Condition
ITEM INFORMATION:       Financial Statements and Exhibits
FILED AS OF DATE:       20200123
DATE AS OF CHANGE:      20200123

FILER:

COMPANY DATA:
COMPANY CONFORMED NAME: AVNET INC
CENTRAL INDEX KEY:      0000008858
STANDARD INDUSTRIAL CLASSIFICATION: WHOLESALE-ELECTRONIC PARTS & EQUIPMENT, NEC [5065]
IRS NUMBER:             111890605
STATE OF INCORPORATION: NY
FISCAL YEAR END:        0627
```

¹ See U.S. Securities and Exchange Commission, “Webmaster Frequently Asked Questions,” available at <https://www.sec.gov/os/webmaster-faq> (last visited Jan. 10, 2023) (noting that “The remaining digits reflect the year of the filing and the sequential count of submitted filings from that CIK.”).

² See <https://www.sec.gov/Archives/edgar/data/8858/000000885820000007/0000008858-20-000007.txt> (last visited Jan. 10, 2023).

In making the filing, however, Avnet used FA 1 software, as EDGAR also makes clear:³

```
</SEC-HEADER>
<DOCUMENT>
<TYPE>8-K
<SEQUENCE>1
<FILENAME>avt-20200123x8k.htm
<DESCRIPTION>8-K
<TEXT>
<XBRL>
<?xml version='1.0' encoding='UTF-8'?>

<!-- iXBRL document created with: REDACTED iXBRL 9.5.7269.28134 -->
<!-- Based on: iXBRL 1.1 -->
<!-- Created on: 1/23/2020 8:54:40 PM -->
<!-- iXBRL Library version: 1.0.7269.28141 -->
<!-- iXBRL Service Job ID: 7ca4df3d-dc62-4f0c-bd64-187085b8a7fa -->
```

Mr. Clarke took account of this fact in his analysis. To determine the frequency of such occurrences, he manually spot checked all of the earnings reports designated as self-filed during the first quarter of 2019 (approximately 2,000 reports), and determined that only two percent of them (41 reports) were attributable to FA 1 software—and none was attributable to FA 2.

Accordingly, Klyushin's contention that the Avnet example undermines Mr. Clarke's statistical analysis is untrue. That is because Mr. Clarke's null hypothesis was that *there was no relationship* between the earnings-related trading by Klyushin (and his alleged co-conspirators) and whether FA 1 or FA 2 handled the issuers' earnings filings.⁴ The analysis showed that, contrary to the null hypothesis, there was a strong correlation between those two things. The fact that a small number of self-filed reports were submitted using FA 1 software is irrelevant to this conclusion. The statistical analysis was unaffected because self-filed reports were neither included in the filings Mr. Clarke attributed to FA 1 or FA 2 (16,810 out of 38,359 total filings), nor were

³ *Id.*

⁴ In statistical analysis, the null hypothesis is the statement being tested: for example, that there is no relationship between two sets of observed data.

they included in the number of earnings events in which Klyushin traded where the filings were handled by FA 1 or FA 2 (343 trades out of 356 earnings events traded). And even if they had been included, the minuscule numbers would have made no difference to Mr. Clarke's conclusions. The self-filed reports do not affect the conclusion that FA 1 and FA 2 handled 44 percent of earnings filings based on accession numbers, or that 96 percent of earnings-related trades by the defendant were attributable (by accession number) to those filing agents, and certainly not materially so.

B. Klyushin's Contention that Data Is Missing Is Unsupported

In a single sentence, Klyushin contends that "there is data missing" from Mr. Clarke's analysis, "including earnings information available for some but not all the transactions." Dkt. 131 at 4. Klyushin cites the following snippet from the data set:

A	B	C	D	E	F	G	H	I
account	company	filed_by_ad_t	anndt	event_number	anndats	earnings	filing_agent	Filename
M13	0001528849	1	9/10/2019 16:15	70614	9/10/2019			edgar/data/1528849/0001558370-19-008556.txt
M13	0001528849	1	12/4/2019 16:23	70615	12/4/2019			edgar/data/1528849/0001558370-19-011335.txt
M13	0001528849	1	3/30/2020 16:05		3/30/2020			edgar/data/1528849/0001558370-20-003383.txt
M13	0001528849	1	6/4/2020 16:05	70617	6/4/2020			edgar/data/1528849/0001558370-20-007197.txt
M13	0001528849	1	9/9/2020 16:16		9/9/2020			edgar/data/1528849/0001558370-20-011091.txt

Though Klyushin does not spell out the basis for his contention that data is missing, he is presumably relying on the absence of entries under the Column G heading "earnings." If so, he misreads the table. *All* of the entries in the table reflect earnings reports. The column titled "earnings" simply reflects earnings reports on which the defendant (or his co-conspirators) traded. Where the column is blank, they did not trade. The example Klyushin cites simply does not support the proposition that data is missing.

C. Klyushin's Remaining Challenges to the Data and Mr. Clarke's Purported Assumptions Are Without Merit

Klyushin alleges several further errors in the relevant data. First, he contends that the market share attributable to FA 1 and FA 2 should be assessed based on the companies' client lists, rather than

the percentage of overall filings they handled over the relevant period. He is mistaken. Just as an individual may be a customer of two car companies—Ford and Toyota—but use only the Ford for trips to the grocery store and only the Toyota trips to the office, the evidence will show that issuers may file their earnings reports in different ways in different quarters. As just one example, between April 2018 and July 2019, for example, Brown & Brown, Inc., a publicly traded insurance company, self-filed its quarterly earnings report using Workiva software. Between October 2019 and July 2020, by contrast, FA 2 filed Brown & Brown’s quarterly earnings reports on its behalf.⁵ The fact that Brown & Brown would appear on FA 2’s client list is misleading for purposes of the analysis in this case, because not all of its filings were handled by FA 2. For this reason, Mr. Clarke’s analysis of individual filings by accession number is a more accurate method of determining market share. While the defense may wish to cross-examine Mr. Clarke about his methodology and conclusions—or seek to offer its own expert’s analysis of market share based on client lists—there is no basis to exclude Mr. Clarke’s measure of market share as unreliable.⁶

Second, Klyushin contends that Mr. Clarke’s data improperly excluded trading in companies unrelated to FA 1 and FA 2 and trading not immediately preceding an earnings announcement. Not so. Mr. Clarke analyzed *all* of the conspirators’ trading between January 2018 and September 2020 and determined that the overwhelming bulk of it was earnings-related—that is, the trades were opened within three days prior to and closed after an earnings announcement. Roughly 98 percent of this

⁵ Notably, Klyushin and his co-conspirators traded around Brown & Brown’s earnings announcements in April 2020 and July 2020, when those announcements were filed by FA 2, but not earlier, when they were self-filed by the company using Workiva software.

⁶ One way of avoiding the dispute altogether may simply be to use a term other than market share, such as “percentage of earnings reports filed during the relevant period.”

trading was in issuers whose earnings announcements were handled by FA 1 and FA 2, and the conspirators collectively earned more than \$97 million from that earnings trading. By contrast, they earned negligible amounts, or lost money, on all of their other trades over the same period, regardless of whether those trades were short-term or long-term, or whether they occurred around the time of earnings announcements.

Only *after* conducting this analysis did Mr. Clarke further analyze Klyushin's earnings-related trading to determine whether there was a correlation between his earnings-related trades and the identity of the filing agent for those earnings announcements. For purposes of this analysis, Mr. Clarke likewise did *not* exclude trading in companies unrelated to FA 1 and FA 2. He found that the probability that the observed trading would have occurred in the absence of a relationship with the identity of the filing agent was less than one in a trillion. In other words, the trading and the identity of the agent were highly correlated.

For all of the defense's various challenges to Mr. Clarke's analysis, the defense does not seriously dispute that this correlation exists, or even that it is statistically significant. Instead, Klyushin contends that there could be other explanations for the correlation.⁷ But Mr. Clarke will not offer any opinion about *why* Klyushin's earnings-related trades are correlated with the identity of the filing agent.

⁷ Klyushin also contends that the large sample size makes it easy to "cherry-pick an arbitrary variable from the earnings events data"—such as whether the digits "32" appear in the Central Index Key assigned to SEC filers—and find this meaningless fact to be statistically significant. Dkt 131-2 at 4. But the fact that Klyushin's expert was able to misuse a widely accepted statistical test to reverse engineer a statistically significant result does not render the proper use of that test unreliable. By Klyushin's reasoning, the ability to manufacture spurious correlations would render all statistical analysis meaningless. The way to distinguish spurious correlations such as those concocted by Klyushin's expert from meaningful correlations is proper test design and common sense. Mr. Clarke's test was properly designed—not reverse engineered—to answer a relevant question: whether there was a statistically significant relationship between observed data. Other evidence the government will introduce at trial will prove *why* that relationship exists.

The government will offer *other* evidence about that, such as chats among the co-conspirators about disguising raw MNPI as the work product of an analyst, and photographs of some of the stolen MNPI on the traders' computer screens prior to its public release. Likewise, Mr. Clarke will *not* testify, as the defense suggests, that his analysis suggests the probability that the data was produced by random chance alone, or that there was a "one-in-a-trillion chance" of Klyushin's trading as he did, or "that the stock market is a game of chance." Dkt. 131 at 1. Nor does his testimony does assume "that a fair, efficient, properly functioning market operates . . . like a coin flip." *Id.* at 6. All this is a red herring. Mr. Clarke will testify, to the contrary, that chance alone *cannot* explain the correlation between Klyushin's trades and the identity of the filing agent. Indeed, such trading would be expected to occur less than one in a trillion times *in the absence* of a relationship. Mathematically—as the Court pointed out—this is not unlike the probability that a coin, flipped 400 times, will turn up heads 96 percent of the time. That is because there is no reason to expect a correlation between an issuer's ministerial choice of filing agent and the defendant's choice of what stock to trade—thus, Mr. Clarke's "null hypothesis" that no such correlation exists.⁸

⁸ Though Klyushin focuses his challenges on Mr. Clarke's analysis of the correlation between the issuers' choice of filing agents and the conspirators' trading, he also makes passing, undeveloped challenges to Mr. Clarke's analysis of the correlation between (1) outsized earnings surprises and the direction of the conspirators' trades, and (2) the time that earnings reports were downloaded from FA 2 and when the conspirators traded. As to the former, Klyushin again suggests that Mr. Clarke will testify "that there is a 'one in a trillion' chance Mr. Klyushin correctly predicted whether to buy or sell a stock," which he contends is premised on the "assumption that traders randomly buy or sell stocks." Dkt. 131 at 6. That is not correct. Mr. Clarke's null hypothesis was that there should be *no* correlation between the *direction* of the conspirators' stock trades (long or short) and the occurrence of large earnings surprises—that is, earnings that exceeded, or fell short of, analysts' expectations by more than a quartile. Once again, he concluded that there was a strong correlation between those variables. Put differently, the observed pattern of trading would have occurred less than one in a trillion times in the absence of a relationship. As to the latter, Klyushin contends that Mr. Clarke "purposely ignore[ed] any transactions, particularly transactions by Sladkov or Irzak, in the same stocks prior to the alleged hack." *Id.* at 7. In fact, neither Irzak nor any of the other traders besides Sladkov had trading accounts prior to

II. Klyushin's Challenge to the Use of Fisher's Exact Test Is Meritless

Klyushin challenges Mr. Clarke's use of Fisher's exact test on the basis of his rebuttal expert's unsupported contention that the test is only appropriate for "simple, controlled experiments," and is "poorly suited" for data "collected from historical events." Dkt. 131 at 8. That is not true. Mr. Clarke employed the test because—where there is sufficient computing power to conduct the analysis—the test is valid for all sample sizes. And because it analyzes *all* possible combinations of outcomes, it permits the precise determination of the probability of an observed relationship between two variables. Here, those are (1) the issuer's choice of filing agent on the one hand, and the conspirators' choice of stocks on the other; and (2) outsized earnings surprises on the one hand, and the direction of the conspirators' trades on the other. Klyushin criticizes the application of the test here insofar as it fails to account for other, potentially explanatory variables. Tellingly, however, his own expert fails to offer any such variables that might explain his trading. Mr. Clarke, too, will not offer any. He will simply testify as to the statistical significance of the observed relationships. If asked, however, he will also testify that he analyzed the same data using *other* statistical tests, including the chi-squared test—which is specifically used for large sample sizes—the T-test and the ANOVA Test, and all yielded substantially the same results.

2018 of which the government is aware. And Klyushin's contention that evidence of prior stock trading by Sladkov would prove that his trades "were not based on alleged intrusions," *id.*, is nonsense. Even assuming, *arguendo*, that Sladkov traded lawfully *prior* to any known intrusion, that is not evidence that he traded lawfully *after* the known intrusion—just as evidence that an individual lawfully used an ATM on one day would not be evidence that he did not rob the bank the next.

III. The Testimony Is Relevant and Probative and Should Not Be Excluded

Failing all else, Klyushin asks the Court to exclude Mr. Clarke’s testimony under Rule 403. The Court should decline this invitation. The testimony is highly relevant and uniquely probative. And its probative value is not substantially outweighed by the danger of unfair prejudice.

As an initial matter, the First Circuit has repeatedly recognized that *all* evidence is prejudicial, and “it is only unfair prejudice against which the law protects.” *United States v. Simon*, 12 F.4th 1, 42 & n.9 (1st Cir. 2021) (cleaned up; emphasis in original) (collecting cases); *see also United States v. Whitney*, 524 F.3d 134, 141 (1st Cir. 2008) (“As Rule 403 explicitly states, the law shields defendant against unfair prejudice, not against all prejudice.”) (cleaned up); *United States v. Rodriguez–Estrada*, 877 F.2d 153, 156 (1st Cir.1989) (“By design, all evidence is meant to be prejudicial; it is only unfair prejudice which must be avoided.”) (cleaned up). To run afoul of Rule 403, “[t]he admitted evidence must not only be prejudicial, but be unfairly prejudicial, and not only outweigh relevance but substantially outweigh relevance.” *United States v. Rivera*, 83 F.3d 542, 545 (1st Cir. 1996). Moreover, “where the reviewing court finds the balancing close, *Rule 403 tilts the balance in favor of admission.*” *Id.* (emphasis added).

Here, for all the reasons noted, Mr. Clarke’s statistical analysis is highly probative. It will, for example, help the jury understand that there is a relationship between Klyushin’s trading and the fact that the issuers in whose securities he traded used FA1 or FA2, while not relieving the government of the need to prove, beyond a reasonable doubt, *why* that correlation exists, using other evidence. Indeed, where, as here, there is a huge volume of trading by multiple traders in hundreds of stocks over an extended period of time, it would be impractical, and require a trial lasting weeks (if not months) to link each individual download with an individual trade. Moreover, Mr. Clarke will, as noted, offer no opinion about *why* the conspirators’ trades correlate to the issuers’ choice of filing agent or the timing

of earnings report downloads, or *why* the direction of their trades correlate to outsized earnings surprises. The government will offer other evidence on these points.

Nor is the degree of the statistical correlation unduly prejudicial. *See, e.g., United States v. Pritchard*, 993 F. Supp. 3d 1203, 1207, 1213–14 (C.D. Cal. 2014) (allowing expert testimony that the probability of a coincidental DNA match was “less than 1 in 1 trillion,” and rejecting argument that jury would accept evidence “as a statement of . . . the likelihood that the defendant is the source of the evidentiary sample . . . [and] equate source with guilt”). In any event, to the extent the Court has concerns about prejudice, it can offer a limiting instruction about how such statistical evidence may be considered—which the First Circuit has made clear is preferable to excluding probative evidence altogether. *See, e.g., United States v. Santana*, 342 F.3d 60, 67 (1st Cir. 2003) (“[P]rejudicial effect can be reduced by issuing an appropriate limiting instruction.”) (cleaned up); *United States v. Escobar-de Jesus*, 187 F.3d 148, 169 n.20 (1st Cir. 1999) (courts “should . . . issue a limiting instruction to the jury concerning the purpose of such evidence in order to further reduce the potential for prejudice flowing from its admission”).

Klyushin contends that Mr. Clarke’s analysis “fails to account for the possibility that the patterns arose from factors other than insider trading based on hacked information, e.g., informational advantages obtained through legitimate means such as effort, analysis, study and skill.” Dkt. 131 at 11. But that is not true. As noted, the analysis itself offers *no* explanation for the patterns. If Klyushin wishes to argue, or offer evidence, that he earned tens of millions of dollars over less than two years by trading stocks associated with just two filing agents based on legitimate analysis, study and skill, he is free to do so. *See, e.g., Pritchard*, 993 F. Supp. 2d at 1214 (“To the extent that there is any possibility of non-criminal reasons for Defendants’ DNA being found on any of the items or surfaces swabbed . . . Defendants’ lawyers are very capable of making that argument to the jury.”). The government will offer evidence that he obtained his informational advantage through other means: by

stealing MNPI from the filing agents and trading on it. Just because the observed correlations may be difficult for Klyushin to plausibly explain, does not mean that the jury should be deprived of evidence that they exist, and are statistically highly significant. Mr. Clarke's analysis is critical to distilling voluminous evidence and helping the jury understand it. There is no basis to exclude it.

Respectfully submitted,

RACHAEL S. ROLLINS
United States Attorney

By: /s/ Stephen E. Frank
STEPHEN E. FRANK
SETH B. KOSTO
Assistant U.S. Attorneys

Date: January 11, 2023

CERTIFICATE OF SERVICE

I hereby certify that a copy of this document will be sent electronically to the registered participants as identified on the Notice of Electronic Filing.

/s/ Stephen E. Frank
STEPHEN E. FRANK
Assistant U.S. Attorney